On the validity and errors of the pseudo-first-order kinetics in ligand–receptor binding

Wylie Stroberg[a], Santiago Schnell[a,b,c,*]

[a] *Department of Molecular & Integrative Physiology, University of Michigan Medical School, Ann Arbor, MI 48109, USA*
[b] *Department of Computational Medicine & Bioinformatics, University of Michigan Medical School, Ann Arbor, MI 48109, USA*
[c] *Brehm Center for Diabetes Research, University of Michigan Medical School, Ann Arbor, MI 48105, USA*

Abstract

The simple bimolecular ligand–receptor binding interaction is often linearized by assuming pseudo-first-order kinetics when one species is present in excess. Here, a phase-plane analysis allows the derivation of a new condition for the validity of pseudo-first-order kinetics that is independent of the initial receptor concentration. The validity of the derived condition is analyzed from two viewpoints. In the first, time courses of the exact and approximate solutions to the ligand–receptor rate equations are compared when all rate constants are known. The second viewpoint assess the validity through the error induced when the approximate equation is used to estimate kinetic constants from data. Although these two interpretations of validity are often assumed to coincide, we show that they are distinct, and that large errors are possible in estimated kinetic constants, even when the linearized and exact rate equations provide nearly identical solutions.

Keywords: pseudo-first-order kinetics, ligand–receptor binding, experimental design, approximation validity, rate constant estimation, fitting procedure.

*Corresponding author.
Email addresses: stroberg@umich.edu (Wylie Stroberg), schnells@umich.edu (Santiago Schnell)

1. Introduction

In biochemical kinetics, simplifying assumptions that decouple or reduce the order of rate equations for complex reaction mechanisms are ubiquitous. Aside from making theoretical analysis of complex reactions more tractable, order-reducing approximations can greatly simplify the interpretation of experimental data [1, 2]. Experiments performed under conditions that allow for linearization have historically been the preferred method for estimating equilibrium and rate constants because they allow for the isolation of a subset of the interactions [3, 4, 5]. For this reason, when designing an experiment, it is essential to know the necessary conditions for the simplifying assumptions to be valid. Significant theoretical work has been directed at deriving rigorous bounds for the validity of simplifying assumptions [6, 7, 8, 9, 10, 11], but this work often overlooks the manner in which the reduced models are used to interpret experimental results. In many cases, the simplified models are used to estimate equilibrium and rate constants from experimental data [12, 13, 14, for example]. Rarely is the validity of a simplifying assumption analyzed with this utility in mind. To examine how this viewpoint can affect the conditions for validity, we consider the simplest model for ligand–receptor binding with 1:1 stoichiometry [15].

In the simplest case, the binding of a ligand L to a receptor R is a bimolecular reversible association reaction with 1:1 stroichiometry yielding a ligand–receptor intermediate complex C:

$$L + R \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C \ , \tag{1}$$

where k_1 and k_{-1} are, respectively, the association and dissociation rate constants of the ligand–receptor complex. This reaction scheme is mathematically described by a system of coupled nonlinear second-order differential equations. By applying the law of mass action to reaction (1), we obtain

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = -\frac{\mathrm{d}[R]}{\mathrm{d}t} = -\frac{\mathrm{d}[L]}{\mathrm{d}t} = k_1\Big([R][L] - K_S[C]\Big) \ . \tag{2}$$

In this system the parameter $K_S = k_{-1}/k_1$ is the equilibrium constant [15, 4] and the square brackets denote concentration. Since no catalytic processes are involved, the reaction is subject to the following conservation laws:

$$[R_0] = [R](t) + [C](t) \tag{3}$$

$$[L_0] = [L](t) + [C](t) \ , \tag{4}$$

where $[R_0]$ and $[L_0]$ are the initial receptor and initial ligand concentrations. If the bimolecular reaction (1) is initiated far from the equilibrium and in the absence of ligand–receptor complex, the system (2) has the initial conditions at $t = 0$:

$$([L], [R], [C]) = ([L_0], [R_0], 0) \ . \tag{5}$$

We have expressed quantities in terms of concentration of species. These equations are frequently given in terms of binding site number, using the identity [15]

$$[C] = \Big(\frac{n}{N_{AV}}\Big)C \ , \tag{6}$$

where n is the cell density, N_{AV} is Avogadro's number, and C denotes the number of ligand-bound receptors per cell. We use the concentration formulation here for clarity and without loss of generality.

The system (2) can be solved, subject to the conservation laws [16]. Substituting (3) and (4) into (2) we obtain:

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_1\Big(([R_0] - [C])([L_0] - [C]) - K_S[C]\Big) \ . \tag{7}$$

We can rewrite this expression by factoring as follows:

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_1\big((\lambda_+ - [C])(\lambda_- - [C])\big) \ , \tag{8}$$

where

$$\lambda_\pm = \frac{(K_S + [R_0] + [L_0]) \pm \sqrt{(K_S + [R_0] + [L_0])^2 - 4[R_0][L_0]}}{2} \ . \tag{9}$$

This ordinary differential equation is readily solved subject to the initial conditions (5) as

$$[C](t) = \lambda_- \left(\frac{1 - \exp(-\frac{t}{t_C})}{1 - \frac{\lambda_-}{\lambda_+}\exp(-\frac{t}{t_C})}\right) \ , \tag{10}$$

with

$$t_C = \left[k_1\sqrt{(K_S + [R_0] + [L_0])^2 - 4[R_0][L_0]}\right]^{-1} \ . \tag{11}$$

The quantity t_C is the timescale for significant change in $[C]$. In this particular case, t_C can be considered as the time required for the reaction to

reach steady-state. Solutions for $[R](t)$ and $[L](t)$ can now be constructed by substituting (10) into conservation laws (3) and (4).

Although there is a closed form solution for the reacting species of the simple bimolecular ligand–receptor interaction, experimental biochemists prefer to determine the kinetic parameters of the ligand–receptor binding using graphical methods [15]. One of the graphical methods commonly used consists of plotting the solution of the ligand association assuming no ligand depletion on a logarithmic scale with respect to time. Both the association and dissociation rate constants can be determined using this linear graphical method [17]. Similarly, if one seeks to avoid inaccuracies due to logarithmic fitting, nonlinear regression can be used to fit the kinetic data to a single exponential. However, the use of both of these methods has the disadvantage of making an assumption with respect to the relative concentrations of ligand and binding sites [16].

In the ligand–receptor interaction with 1:1 stochiometry and no ligand depletion it is generally thought that, if the initial ligand concentration is much higher than the initial receptor concentration, i.e.

$$[L_0] \gg [R_0] \ , \tag{12}$$

the ligand concentration $[L]$ remains effectively constant during the course of the reaction, and only the receptor concentration $[R]$ changes appreciably with time [18, 19, 3, 4]. Since kinetic order with respect to time is the same as with respect to $[R]$, reaction (1) is said to follow *pseudo-first-order* (PFO) kinetics if the $[R]$ dependence is of first order. The rates of second-order reactions in chemistry are frequently studied within PFO kinetics [20, 21]. In the present case, the second-order reaction (1) becomes mathematically equivalent to a first-order reaction, reducing to

$$R \underset{k_{-1}}{\overset{k_\varphi}{\rightleftharpoons}} C \ , \tag{13}$$

where $k_\varphi \equiv k_1[L_0]$ is the pseudo rate constant. This procedure is also known as the method of flooding [5]. The solution of the governing equations for a reaction linearized by PFO kinetics (or flooding) is straightforward, and is widely employed to characterize kinetics and fit parameters with the aid of progress curves. An error is however present due to the fact that, in actuality, the concentration of the excess reactant does not remain constant [20].

In 1961, Silicio and Peterson [20] made numerical estimates for the fractional error in the observed PFO constant for second-order reactions. They found that the the fractional error is less than 10% if the reactant concentration ratio, $[R_0] : [L_0]$ say, is tenfold. On the other hand, Corbett [22] found that simplified expressions with the PFO kinetics can yield more accurate data than is generally realized, even if only a twofold excess of one the reactants is employed. For ligand–receptor dynamics, Weiland and Molinoff [16] claim that the PFO simplification is acceptable if experimental conditions are such that less than 10% of the ligand is bound. These results indicate that the conditions whereby a second-order ligand–receptor reaction is reduced to first order remain to be well-established.

It is widely believed that second-order reactions can be studied by PFO kinetics using progress curves only when the excess concentration of one of the reactants is large [21, 5, for example]. However, contrary to the widely established knowledge, Schnell and Mendoza [10] have found that the condition for the validity of the PFO in the single substrate, single enzyme reaction does not require an excess concentration of one of the reactant with respect to the other. In the present work, we derive the conditions for the validity of the PFO approximation in the simple ligand–receptor interaction. Additionally, we show two fundamentally different methods of assessing the validity of the approximation. The first compares the exact and approximate solutions to the rate equations under identical conditions. The second measures the veracity of parameters estimated by fitting the approximate model to data. Although these two measures of validity are generally assumed to coincide, we show that they are quantitatively and qualitatively distinct. In Section 2 the reduction of the ligand–receptor association by PFO kinetics is summarized followed by its dynamical analysis in Section 3. The new validity condition is derived in Section 4, and an analysis of the errors observed with the PFO kinetics is presented in Section 5. This is followed by a brief discussion (Section 6).

2. The governing equations of the ligand–receptor dynamics with no ligand depletion

In ligand–receptor dynamics with 1:1 stochiometry and no ligand depletion, the second-order ligand–receptor interaction in reaction (1) is neglected when condition (12) holds; the reaction effectively becomes first order since the concentration of the reactant in excess is negligibly affected. This is

equivalent to assuming that

$$[L_0] - [C] \approx [L_0] \ . \tag{14}$$

The alternative case, in which the depletion of the receptor is assumed to be negligible, is shown to be symmetric in Appendix A. By substituting (14) into (7), the equation can be simplified as follows:

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_1\Big([R_0][L_0] - (K_S + [L_0])[C]\Big) \ . \tag{15}$$

Note that this equation can also be obtained by applying the law of mass action to reaction scheme (13). The solution for (15) with the conservation laws (4) and (5) can be obtained by direct integration [23]:

$$[C](t) = \frac{[R_0][L_0]}{K_S + [L_0]}\Big(1 - \exp\left(-(k_\varphi + k_{-1})t\right)\Big) \ . \tag{16}$$

Note that expressions for $[R](t)$ and $[L](t)$ can again be obtained by substituting (16) into (3) and (4), respectively.

Experimentally, the clear advantage of applying the pseudo-first-order kinetics to the ligand–receptor reaction is that, as shown by equation (16), it provides solutions that can be linearized by using a logarithmic scale to fit progress curves of the interacting species and thus it could lead to complete reaction characterization, namely the rate constants k_1 and k_{-1}.

As we have previously pointed out, it has been assumed that the condition $[L_0] \gg [R_0]$ implies

$$[L](t) = [L_0] - [C](t) \approx [L_0] \qquad \Rightarrow \qquad \left|\frac{[C]}{[L_0]}\right|_{\max} \ll 1 \ . \tag{17}$$

Up to this point, most of the scientists using the PFO kinetics assume that it is reasonable to overestimate the maximum complex concentration when there is a ligand excess, because all receptor molecules could instantaneously combine with ligand molecules, i.e. $[C]_{\max} = [R_0]$. However, this simplification is unrealistic from the biophysical chemistry point of view as it will assume that in the conservation law (4) all ligand molecules are only in one form: the free ligand [see, equation (17)]. In the next section, we obtain a more a reliable estimate of $[C]_{\max}$ by studying the geometry of the phase plane of system (2). This will permit us to make a better estimate of the conditions for the validity of the PFO ligand–receptor dynamics (1).

3. Phase-plane analysis leads to conditions for the validity of the pseudo-first-order kinetics

The phase-plane trajectories of system (2) are determined by the ratio of $\mathrm{d}[C]/\mathrm{d}t$ to $\mathrm{d}[L]/\mathrm{d}t$:

$$\frac{\mathrm{d}[C]}{\mathrm{d}[L]} = -1 \ . \tag{18}$$

This expression is integrated to obtain the family of solution curves:

$$[C]([L]) = -[L] + m \ , \tag{19}$$

where

$$m = \begin{cases} [C_0] & \text{for } ([C],[L]) = (0,[L_0]) \text{ at } t = 0 \\ [L_0] & \text{for } ([C],[L]) = ([C_0],0) \text{ at } t = 0 \ . \end{cases}$$

The use of $[L_0]$ as the constant in (19) for initial conditions on the horizontal axis follows from the relation $\mathrm{d}[C]/\mathrm{d}t = -\mathrm{d}[L]/\mathrm{d}t$.

The phase plane is divided into two regions by the nullcline

$$[C]([L]) = \frac{[R_0][L]}{K_S + [L]} \ , \tag{20}$$

obtained by setting $\mathrm{d}[C]/\mathrm{d}t = 0$ or $\mathrm{d}[L]/\mathrm{d}t = 0$ in (2). Note that for $[L] = K_S$ (equilibrium constant), $[C] = [R_0]/2$. The nullcline converges at large ligand concentrations because

$$\lim_{[L]\to\infty} [C]([L]) = [R_0] \ . \tag{21}$$

The phase plane trajectories (19) and its nullcline (20) are show in Fig. 1. The trajectory flow is attracted by a unique curve, which is a stable manifold and is equivalent to the nullcline for this case. All trajectories tend to this manifold as they approach the steady state as $t \to \infty$ [24].

Binding of ligand to cell surface receptors has been amenable to *in vitro* experimental investigation for the past four decades [25]. In the typical experimental approach, isolated membranes possessing free receptors are studied using ligands as pharmaceutical agents [26]. The reaction mixture is free of ligand–receptor complex at the beginning of the experiment, that is the initial conditions are like those stated in (5). It is important to note that for a ligand-receptor interaction with trajectories departing from the positive

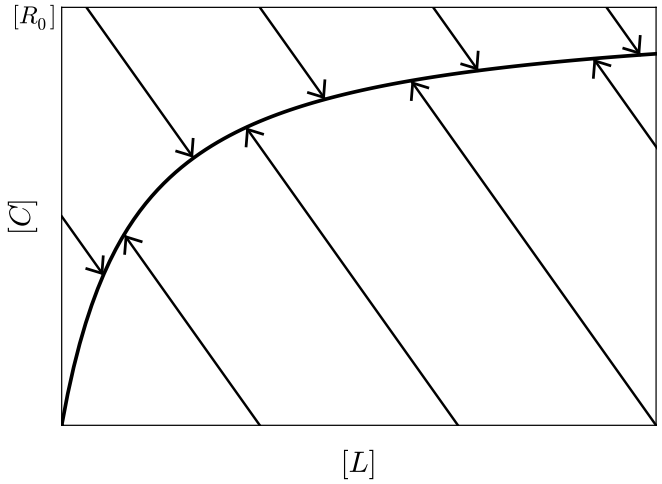

Figure 1: **Phase-plane behavior of the ligand–receptor reaction (1)**. The solid curves with arrows are the trajectories in the phase planes, which are described by (19). The trajectories tend to a stable manifold as they approach to the steady-state. In this case, the manifold is the nullcline (20) of the system, which converges to $[R_0]$ for large ligand concentrations.

horizontal axis, i.e. with initial conditions (5), the trajectories are bounded by

$$0 \leq [C](L) \leq [C]^\star \qquad \text{for} \quad ([L],[C])(t=0) = ([L_0],0) \ , \tag{22}$$

where $[C]^\star$ is the ligand–receptor complex concentration at the steady-state, and is equivalent to the maximum ligand–receptor complex concentration ($[C]_{\max}$) that the trajectories can reach if they depart from the positive horizontal axis. $[C]^\star$ can be estimated from the intersection of (19) and (20) or by estimating the steady-state value of the ligand–receptor complex concentration in (10), that is

$$[C]^\star = \lim_{t\to\infty}[C](t) = \lim_{t\to\infty}\left[\lambda_-\left(\frac{1-\exp(-\frac{t}{t_C})}{1-\frac{\lambda_-}{\lambda_+}\exp(-\frac{t}{t_C})}\right)\right] = \lambda_- \ , \tag{23}$$

where λ_- is given by (9).

It suffices therefore to investigate the behavior of the ratio of the solution (10) at steady-state to $[L_0]$, which we do in the next section.

4. Derivation of a new sufficient condition for the validity of the pseudo-first-order kinetics

To derive a mathematical expression in terms of the kinetic parameters for condition (17), we will use the fact that, for initial conditions given by (22), $[C]_{max}$ is the concentration given by allowing the reaction described by (10) to go to steady-state. We can now formulate (17) as follows:

$$\frac{(K_S + [R_0] + [L_0]) - \sqrt{(K_S + [R_0] + [L_0])^2 - 4[R_0][L_0]}}{2[L_0]} \ll 1 \ . \tag{24}$$

This can be rewritten as

$$\left(\frac{(K_S + [R_0] + [L_0])}{2[L_0]}\right)\left(1 - \sqrt{1-r}\right) \ll 1 \ , \tag{25}$$

with

$$r = \frac{4[R_0][L_0]}{(K_S + [R_0] + [L_0])^2} \ .$$

At this point it is convenient to nondimensionalize the above expression by using reduced concentrations. Scaling with respect to K_S, equation (25) becomes

$$\left(\frac{(1 + [R_0'] + [L_0'])}{2[L_0']}\right)\left(1 - \sqrt{1-r'}\right) \ll 1 \ , \tag{26}$$

where

$$r' = \frac{4[R_0'][L_0']}{(1 + [R_0'] + [L_0'])^2} \ , \quad \text{with} \quad [R_0'] = \frac{[R_0]}{K_S} \quad \text{and} \quad [L_0'] = \frac{[L_0]}{K_S} \ . \tag{27}$$

Quadratic expressions similar to (26) are common in chemical kinetics. For practical use in the analysis of chemical kinetics experiments, quadratic expressions are generally replaced with simpler expressions. Noting that

$$r' = \frac{4[R_0'][L_0']}{(1 + [R_0'] + [L_0'])^2} \ll 1 \tag{28}$$

for any value of $[R_0']$ and $[L_0']$ (for more details, see Appendix B), we can then calculate a Taylor series expansion of (26) to obtain right-hand factor of

$$\left(1 - \sqrt{1-r'}\right) = \frac{1}{2}r' + \frac{1}{8}r'^2 + O(r'^3) \approx \frac{1}{2}r' \ , \tag{29}$$

which simplifies (26) to

$$\frac{[R_0']}{1+[R_0']+[L_0']} \ll 1 \ . \tag{30}$$

This is a simple analytical expression for the condition for the validity of PFO kinetics. Note that condition (30) is valid for

$$\begin{aligned} &\text{(a)} \qquad [L_0'] \gg [R_0'] \ , \ or \\ &\text{(b)} \qquad [L_0'] \leq [R_0'] \ , \text{ if } [R_0'] \ll 1 \ . \end{aligned} \tag{31}$$

Interestingly, we have a new condition for the use of the PFO approximation in ligand–receptor binding with 1:1 stoichiometry. Equation (31a) is the constraint (12) already in widespread use. However, equation (31b) extends the range of conditions under which PFO dynamics may be applied. The regions of validity of the PFO approximation are illustrated graphically in Fig. 2 by plotting conditions (30) and (31) in the space of initial ligand concentrations, $[L_0]$, and equilibrium constant, K_S, normalized by $[R_0]$. Typically, PFO kinetics are assumed valid when the ratio of $[L_0]$:$[R_0]$ is greater than 10:1 [20, 3, 27] Applying this same "rule of thumb", we set the left-hand side of (30) equal to 0.1 to separate valid from non-valid regions in Fig. 2. Note that the plane is divided into three regions by lines corresponding to condition (30) and the generally used condition (12). Region B comprises a portion of the space where PFO kinetics were previously assumed to be invalid, but in fact the errors introduced by the approximation in this region are expected to be small, even for initial conditions such that $[L_0] \approx [R_0]$.

5. There are two types of errors observed with the application of approximations in reaction kinetics

In reaction kinetics, there are two type of errors that can be committed when applying an approximation to the governing equations of a complex reaction. Research in mathematical chemistry and biology primarily focuses on estimates of errors in the concentrations of reacting and product species. This error – which we name *concentration error* – is commonly evaluated by calculating the difference between the solution of the approximate equation (16) with that of the exact equation (10) (see, for example, [28]). The concentration error provides a measure of how closely the approximate solution matches the exact solution. However, experimentally, PFO approximations are often used to derive expressions that facilitate estimating kinetic

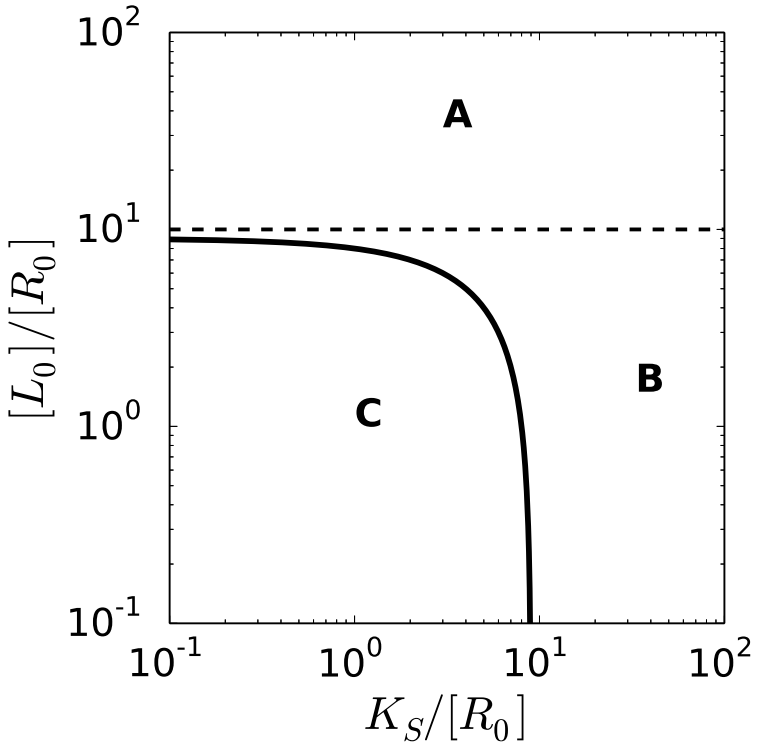

Figure 2: **Validity regions in the $[L_0]/[R_0]$–$K_S/[R_0]$ log-log plane for the use of pseudo-first-order kinetics to model ligand–receptor reaction (1).** The dashed line indicates $10[R_0] = [L_0]$, the lower line is $9[R_0] = [L_0] + K_S$. In region A, $[R_0] \gg [L_0]$, where PFO kinetics has here been shown to be valid, as previously thought. In region B, $[R_0]$ does not greatly exceed $[L_0]$, but $K_S \gg [R_0]$. Here we have shown that PFO kinetics holds, even for $[L_0] \approx [R_0]$. In region C, where $[R_0]$ does not greatly exceed $[L_0]$ and $K_S/[R_0]$ is not much greater than 1, PFO kinetics is not valid.

constants using nonlinear regression methods. Therefore, of particular utility are error estimates from fitting kinetic parameters using the mathematical expression derived with the PFO approximation. We called these *estimation errors*.

Naively, it would seem that if the difference between the complex concentration, $C(t)$, is small between the PFO approximation and exact solution, then the PFO equation should provide accurate estimates of the rate con-

stants when used to model experimental data. This, however, is not necessarily true in general. To better understand why the concentration and estimation errors are not the same, and to show where they diverge, we compare errors based on the numerical difference between the exact and PFO models with those based on estimating rates constants using the PFO model.

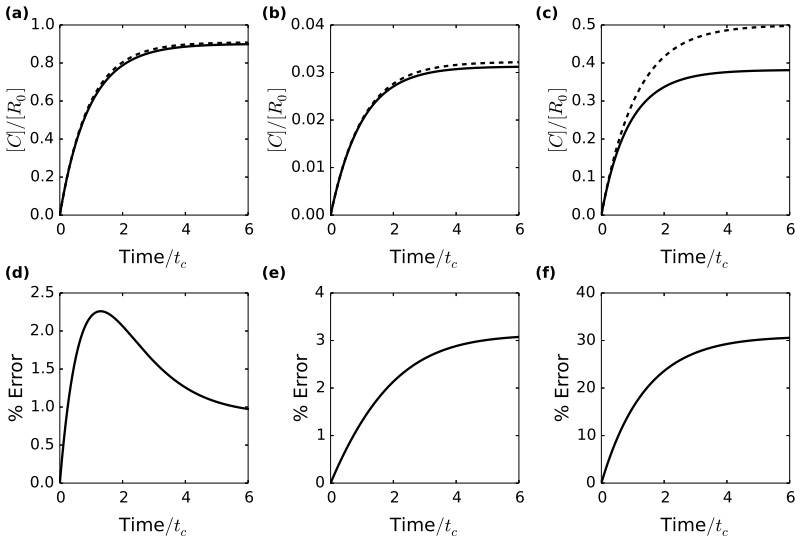

Figure 3: **Time course of species concentrations and concentration errors of pseudo-first-order approximation for the ligand–receptor reaction (1).** Panels (a)-(c) show the concentration of the complex as a function of time for the cases: (a) $[L_0] = 10[R_0]$ and $K_S = [R_0]$, (b) $[L_0] = [R_0]$ and $K_S = 30[R_0]$, (c) $[L_0] = [R_0]$ and $K_S = [R_0]$. The dashed lines correspond to calculations assuming pseudo-first-order kinetics, while solid lines are exact solutions. (d)-(f) show the errors induced by assuming pseudo-first-order kinetics for case (a)-(c), respectively.

5.1. Analysis of the concentration error

Theoretically we define a concentration error measure as

$$\mathrm{CE}(t) = \left| \frac{[C]_{\mathrm{exact}}(t) - [C]_{\mathrm{PFO}}(t)}{C_{\mathrm{exact}}(t)} \right|. \tag{32}$$

For the bimolecular ligand–receptor binding (1), we can calculate analytically the above expression by replacing $[C]_{\mathrm{exact}}(t)$ with (10), and $[C]_{\mathrm{PFO}}(t)$ with (16). However, this expression is too cumbersome, and hence we present a

numerical analysis of the concentration error. Fig. 3 presents the time course of the exact and approximate complex concentration [Fig. 3(a)–(c)], and a calculation of the percentage concentration error ($100 \times \mathrm{CE}$) [Fig. 3(d)–(f)] introduced by the PFO approximation for initial conditions lying in region A, B and C of Fig. 2, respectively. The error remains less than 3% over the course of the reaction for the cases satisfying condition (30), and approaches 30% for the point in region C.

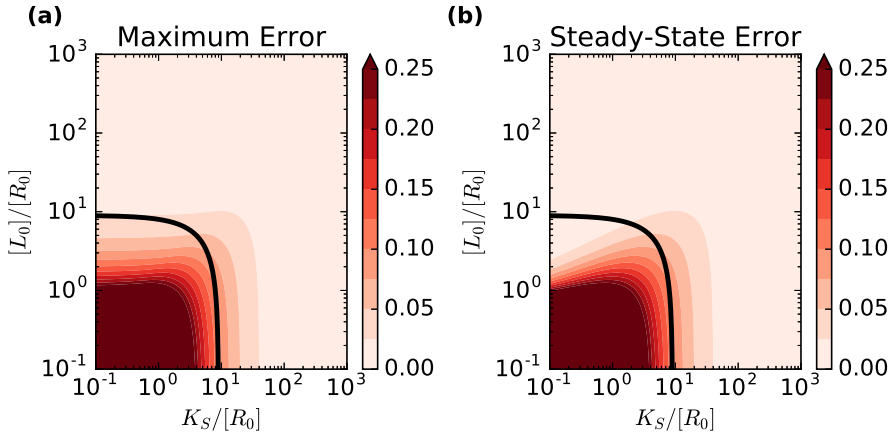

Figure 4: **Maximum and steady-state concentration errors for the ligand–receptor reaction (1).** Panel (a) shows a heat map of the maximum concentration error incurred by assuming pseudo-first-order approximation for different initial conditions. Coloring corresponds to the error as defined in (32). Similarly, Panel (b) shows the steady-state concentration error between the exact and pseudo-first-order solutions. The black lines correspond to condition (30) when the left-hand side is equal to 10.

It is useful to define a scalar measure based on (32). For this there are many options, yet in order to remain as conservative as possible, we choose the maximum value of concentration error over the time course of the reaction, which we call the *maximum concentration error*. Additionally, we calculate the *steady-state concentration error*, defined as $\lim_{t\to\infty} \mathrm{CE}\,(t)$. Fig. 4 shows the maximum and steady-state concentration error contours for initial conditions in the $[L_0]$–K_S plane. In general, the maximum concentration error is well-described by the newly-derived condition. The error contours allow for a quantification of what is meant by "much less than". The commonly used requirement that $[L_0] \geq 10[R_0]$ produces errors always

less than 5%. In fact, when K_S is less than $[R_0]$, a ratio of ligand to receptor of approximately 4:1 is enough to constrain the error to below 5%. Although the condition for the validity is symmetric in $[L_0]$ and $[K_S]$, the errors are not. The ratio of K_S:$[R_0]$ must be approximately 20:1 for the error to remain below 5%. This is not surprising since the exact solution is not symmetric in $[L_0]$ and K_S, so we should expect that the two quantities would have different effects. Nevertheless, the notion that PFO kinetics can rightly be assumed even when the ligand is not present in excess holds true.

5.2. Analysis of the estimation error

Next, we calculate the error in estimated rate constants by generating sample data using the exact solution. The frequency of the sampling, ω_{obs}, and the time span of the sampling window, t_{obs}, are varied. Values of ω_{obs} range from t_c^{-1} to $4t_c^{-1}$. Higher sampling frequencies were also tested, but results are not presented as they did not show appreciable difference from the case of $\omega_{obs} = 4t_c^{-1}$. The sampling windows tested begin at $t = 0$ and continue for $t_{obs} = 3t_c$, $10t_c$, $100t_c$. An "experimental protocol" for a numerical experiment then consists of choosing specific ω_{obs} and t_{obs}, and using (10) to calculate $[C]\left(n\omega_{obs}^{-1}\right)$ for integers $n \in [0, t_{obs}\omega_{obs}]$. For each simulated data set (for which there is no experimental error), the rate constants k_1 and k_{-1} are estimated by fitting the data with equation (16) using the Levenberg–Marquardt algorithm as implemented in SciPy (version 0.17.0, http://www.scipy.org) with initial estimates for k_1 and k_{-1} equal to the values used to generate the data. We then define the *estimation error* of a parameter k_i as

$$\mathrm{EE}\left(k_i\right) = \left|\frac{k_i - k_i^\star}{k_i}\right|, \tag{33}$$

where $k_i^\star$ is the estimate of k_i calculated from fitting the PFO solution to the generated data. Additionally, for the ligand–receptor interaction, we calculate the *mean estimation error* as an aggregate measure of the parameter estimation

$$\mathrm{MEE} = \mathrm{mean}\left\{\mathrm{EE}\left(k_1\right), \mathrm{EE}\left(k_{-1}\right)\right\}. \tag{34}$$

Concentration error measures, such as the maximum or steady-state concentration errors, that compare exact and approximate solutions to the ligand–receptor complex concentration, are fundamentally different from those incurred by using an approximate model to fit experimental data. To illustrate this point, Fig. 5 shows contours for the mean estimation error when different

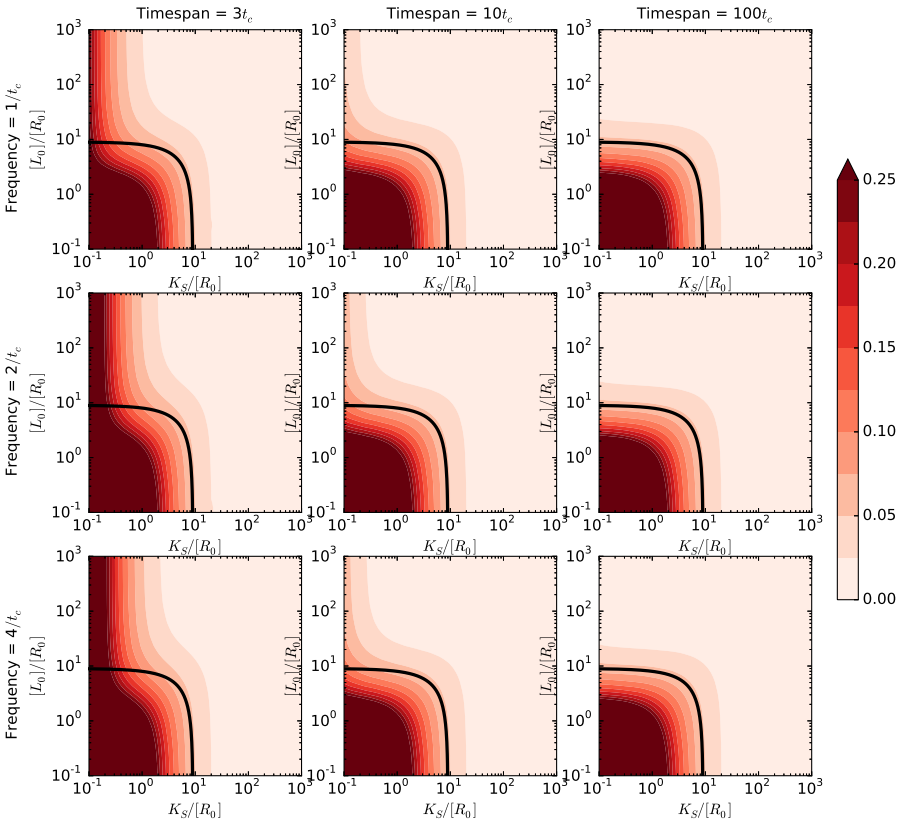

Figure 5: **Mean estimation error in the rate constants when applying the pseudo-first-order approximation to the ligand–receptor reaction (1).** Heat maps of the mean estimation error of the rate constants are shown for different experimental protocols. Columns, from left to right, correspond to increasing length of observation time. Rows, from top to bottom, correspond to increasing frequency of sampling. For short observation windows, large errors occur even when $[L_0] \gg [R_0]$. Also, counter-intuitively, the error at some initial conditions increases as the sampling frequency increases (e.g. down column 1). The black lines correspond to condition (30) when the left-hand side is equal to 10.

"experimental protocols" are used to generate data. For a given sampling frequency, the mean estimation error contours increasing conform to the contour derived from condition (30) as t_{obs} increases. However, for initial conditions

with $[L_0] \geq 10[R_0]$ and $K_S \leq [R_0]$, significant estimation errors can occur if the observation time is not sufficiently long. Even after $10t_c$ of observation, at which point $[C](t) > 0.999[C]^*$, the mean error in the estimated parameters can exceed 10% when K_S is small. Only after nearly $100t_c$ the mean estimation error contours closely mimic the theoretical condition for the validity of the PFO kinetics. This highlights the difference between the errors calculated by comparing the exact and approximate solutions of the concentration equations, and those errors due to fitting an approximate model to data. Additionally, it should caution experimentalists from applying PFO approximations whenever one species is in excess. The values of the rate constants must be considered as well.

Interestingly, increasing the frequency of sampling does not necessarily improve the estimation of the rate constants. In fact, the first column in Fig. 5 shows, that the mean estimation error actually increases as more sample points are used. This effect saturates quickly as the frequency is increased, but nevertheless, using fewer *exact* data points can lead to improved predictions. One major benefit of numerous data points is that it reduces error due to measurement noise, and this likely will outweigh the gains from using fewer data points when fitting. Yet, in cases where accurate measurements are possible, fitting more data to an approximate model can have deleterious effects on the accuracy of parameter estimates made from such a fit. It may be possible to take advantage of both of these effects by recording data at a high frequency, say using an optical assay [29], then performing a number of fits on subsets of the data sampled at lower frequency, thereby reducing both the experimental noise and the estimation error incurred by fitting to an approximate model.

5.3. The error in the estimated parameters for high ligand concentration is due to inaccuracies in k_{-1}

Greater understanding of the estimation errors found at high ligand concentrations and low K_S can be gained by comparing the estimation errors of k_1 and k_{-1} individually. Fig. 6 shows contours of $\mathrm{EE}\,(k_1)$ [panel (a)] and $\mathrm{EE}\,(k_{-1})$ [panel (b)] for numerical data generated with $\omega_{obs} = 2t_c^{-1}$ and $t_{obs} = 3t_c$. From this, it is clear that the inaccuracies lie in predictions of the dissociation rate constant k_{-1}. The reason k_{-1} errors are large for cases in which $[L_0] \gg [R_0]$ and $K_S \ll [R_0]$ can be understood by examining the PFO solution (16) in the limit $K_S/[L_0] \to 0$. Keeping terms up to linear order in

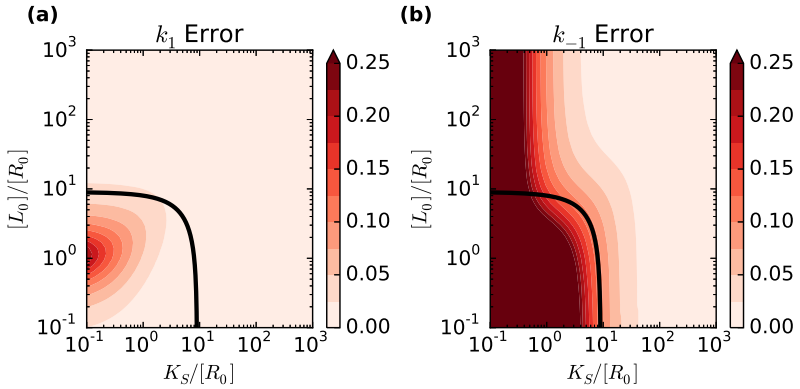

Figure 6: **Comparison of estimation errors for k_1 and k_{-1}.** Heat maps of the EE (k_1) and EE (k_{-1}) are presented in (a) and (b), respectively. The sampling frequency used to generate data was $\omega_{obs} = 2t_c^{-1}$, and the observation window was $t_{obs} = 3t_c$. The black lines correspond to condition (30) when the left-hand side is equal to 10. The error contour for k_{-1} estimation shows clear deviations from the analytical conditions, whereas k_1 estimations are accurate where PFO kinetics is shown to be theoretically valid.

$K_S/[L_0]$, we obtain

$$[C](t) \approx [R_0]\left(1 - \exp\left(-k_1[L_0]t\right)\right)\left[1 - \frac{K_S}{[L_0]}\left(1 - \frac{k_1[L_0]t}{\exp\left(k_1[L_0]t\right) - 1}\right)\right]. \tag{35}$$

The zeroth-order term has no dependence on k_{-1}. Hence, in this limit there is no unique mapping between the parameters $([R_0], [L_0], k_1, k_{-1})$ and time onto the concentration $[C]$, since $[C]$ is completely described by the parameters $([R_0], [L_0], k_1)$ and time. This implies that estimates of k_{-1} from progress curve experiments performed under conditions of high ligand concentration and low disassociation constant are unreliable. However, estimates of the association rate constant k_1 from such experiments should be valid. Unfortunately, since K_S is not generally known a priori, a different measurement, such as an equilibrium binding assay, is required to estimate its value. Then, with knowledge of K_S, k_{-1} can be calculated.

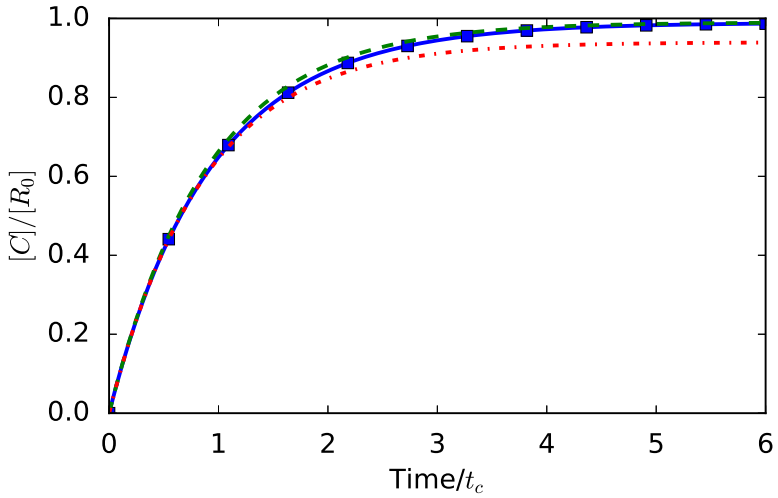

Figure 7: **Comparison of approximate model with exact and fitted parameters for the ligand–receptor reaction (1).** The blue line represents the exact solution and blue squares are simulated data points. The green dashed line is the pseudo-first-order approximation using the same rate constants used to in the exact solution. The red dot-dashed line is the pseudo-first-order approximation using rate constants found by fitting the pseudo-first-order model to the data generated using the exact solution.

5.4. The pseudo-first-order kinetics can lead to significant estimation errors when the conditions of the pseudo-first-order approximation are valid

Taken together, Fig. 4 and Fig. 5 illustrate an important, yet often overlooked, distinction between methods by which to assess the validity of an approximation in chemical kinetics. The first method, popular among theorist, attempts to answer the following question: Given a set of known parameters, how well does the approximate model represent the exact model? This comparison can be made by calculating a maximum or steady-state concentration error, as we do here, or through other measures such as a mean-squared difference over the time course. The second method, which is of greatest importance to experimentalists, answers a subtly different question: Given data, how well do parameters estimated by fitting the data with an approximate model represent the actual parameters? This distinction, between a forward and an inverse problem, is not generally considered when deriving conditions for the validity of reduced kinetic models [30]. To illustrate this, Fig. 7 shows

data points generated with exact solution to the governing equations of the ligand–receptor reaction (1), the PFO solution using the same rate constants as those used to generate the exact data points, and the PFO solution using rate constants calculated through nonlinear regression of the simulated data. The PFO solution using the exact rate constants captures the kinetics much more closely than the PFO solution using estimated rate constants, especially as steady-state is approached. Often, it is the former case that is used by theorists to determine valid ranges for an approximation, while the latter case is where the approximation is actually used to interpret data. As we have shown, these two cases are distinct. Hence, when providing ranges for the validity of a simplifying approximations to theory, it is crucial that the application of that theory be kept in mind.

6. Discussion

We have investigated the application of the PFO approximation to ligand–receptor binding dynamics. PFO kinetics are used to linearize the solutions to the differential equations that describe the concentration of ligand, receptor and ligand–receptor complex over time, allowing them to be fit by a single exponential [16, 15]. This approximation is known to introduce errors that are acceptably small under certain conditions, which have generally been described by $[L_0] \gg [R_0]$. In this paper, we show that this condition is somewhat more stringent than necessary, specifically when $[R_0] \ll K_s$. In fact the condition $[R_0] \ll K_s$ provides another sufficient condition under which one may safely use the PFO approximation, with little regard to the concentration $[L_0]$.

Although it is possible to derive closed-form solutions that describe the kinetics of simple ligand–receptor binding [see, equation (10)], this equation is cumbersome. The PFO approximation gives a much simpler solution [see, equation (16)], which can be linearized by use of logarithmic plots to facilitate data fitting [17]. With this simpler form a linear fit suffices to determine all of the relevant rate constants leading to a complete description of the reaction kinetics. The new condition developed here extends the validity of this method into new territory, increasing its usefulness. Specifically, in cases where reagents are either expensive or difficult to isolate in significant quantities, the new condition suggests far more economical usage (in some cases, orders of magnitude lower concentration) of reagents is possible.

Additionally, we have shown that there is often an inconsistency between the derivation of conditions for validity of an approximation, and the relevant measure of error for applications of that approximation. Approximations to a theory are generally taken as valid if, using identical input parameters, exact and approximate solutions for species concentration are sufficiently similar. This requires minimizing the concentration error introduced in equation (32). However, approximate models are often used to estimate kinetic parameters through fitting to experimental data. We have demonstrated that estimation errors may be significant, even for conditions in which the approximate and exact models are nearly identical. This effect is particularly apparent when one reactant is in excess, the disassociation constant is small, and the length of the observation is not sufficiently long.

Commonly, experimental protocols for kinetic binding assays call for measurements to be made until concentrations "plateau" [27]. However, the definition of "plateau" is often left to the judgment of the investigator. Our analysis shows that stopping measurement prematurely can lead to significant errors in the rate constants estimated by such experiments. A more rigorous definition of the necessary experimental time to reach plateau should involve the inherent timescale, t_c. The error contours in Fig. 5 show that for many initial conditions, specifically for those with large enough K_S, measurements over $3t_c$ are sufficient. It should be possible, however, to test if the experimental conditions are problematic without prior knowledge of K_S, so long as the initial concentrations of receptor is known. If, at steady-state, $[C]$ is very near $[R_0]$, then the affinity of the ligand for the receptor is high enough (and K_S will be small enough) to make the value of k_{-1} from regression analysis unreliable. Since affinities between ligands and receptors are typically quite high, this will often be the case. Hence this suggest that it is necessary to estimate the equilibrium constant using a separate assay. Then, with knowledge of K_S, the rate constants can be unambiguously estimated from kinetic data.

Lastly, we emphasize that the difference between the concentration error and estimation error are not specific to the case of ligand–receptor binding. The problem of estimating parameter values for models is well known in the model calibration field [31], and has also received attention from mathematical and systems biologists [30, 32]. The essential questions are whether or not a parameter in a model actually corresponds to the underlying physical property it is meant to represent, and whether the value of the parameter can be uniquely determined from data. Frequently, the value of the parameter

that provides the best fit to data differs from the most accurate assessment of the underlying physical property, estimated through some other means. In the case of ligand–receptor binding, the rate constants estimated using PFO kinetics correspond to a best-fit of experimental data to an approximate model. In many cases, these estimates will not coincide with the actual rate constants for the second-order reaction. In fact, this difference is quite general and future studies should investigate how the validity of approximations in, for example, Michaelis–Menten kinetics, or inhibited ligand–receptor binding might change when their ability to accurately predict parameters from data is considered.

Appendix A. Symmetry of case with no receptor depletion

The second case referred to in the text (Section 2) applies when the concentration of the receptor $[R_0]$ is much greater than that of the ligand $[L_0]$, which implies that

$$[R_0] - [C] \approx [R_0] \ . \tag{A.1}$$

Substituting (A.1) into equation (7) gives

$$\frac{\mathrm{d}[C]}{\mathrm{d}t} = k_1\Big([R_0][L_0] - (K_S + [R_0])[C]\Big) \ . \tag{A.2}$$

[Compare with (15)]. Solving this equation yields

$$[C](t) = \frac{[R_0][L_0]}{K_S + [R_0]}\Big(1 - \exp(-(k_\varphi + k_{-1})t)\Big) \ . \tag{A.3}$$

This solution is symmetrical with (16). Condition (A.1) gives us the following implication parallel with (17)

$$[R](t) = [R_0] - [C](t) \approx [R_0] \qquad \Rightarrow \qquad \left|\frac{[C]}{[R_0]}\right|_{\max} \ll 1 \ . \tag{A.4}$$

Similar to the case with no ligand depletion, the maximum concentration is equal to λ_-. Hence, following the same procedure as in Section 4, a condition, symmetric to (30), for the case with negligible receptor depletion is found to be

$$\frac{[L_0']}{1 + [L_0'] + [R_0']} \ll 1 \ . \tag{A.5}$$

Appendix B. Validity of approximation (28)

The derivation of the conditions given by (30) requires that equation (28) be satisfied, which we reiterate as

$$r' \ \ll \ 1 \tag{B.1}$$

$$\frac{4[R_0'][L_0']}{(1 + [R_0'] + [L_0'])^2} \ \ll \ 1 \ . \tag{B.2}$$

The above inequality can be written as

$$\frac{1}{1 + [R_0'] + [L_0']} \ll \frac{1}{2\sqrt{[R_0'][L_0']}}. \tag{B.3}$$

Since the denominators are both positive, we can rearrange this as

$$-\left([R_0'] - 2\sqrt{[R_0'][L_0']} + [L_0']\right) \ll 1 \tag{B.4}$$

and factoring the left side then gives

$$-\left(\sqrt{[R_0']} - \sqrt{[L_0']}\right)^2 \ll 1. \tag{B.5}$$

In the above inequalities, the left side is always negative and the right side is clearly positive. Therefore, it is appropriate to assume that $r' \ll 1$.

Acknowledgments

This work is supported by the University of Michigan Protein Folding Diseases Initiative.

References

[1] J. F. Griffiths, Reduced kinetic-models and their application to practical combustion systems, Prog. Energy Combust. Sci. 21 (1995) 25–107.

[2] M. S. Okino, M. L. Mavrovouniotis, Simplification of mathematical models of chemical reaction systems, Chem. Rev. 98 (1998) 391–408.

[3] H. Gutfreund, Kinetics for the life sciences: receptors, transmitters, and catalysts, Cambridge University Press, Cambridge; New York, 1995.

[4] A. Fersht, Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, Macmillan, 1999.

[5] J. H. Espenson, Chemical kinetics and reaction mechanisms, McGraw-Hill : Primis Custom, New York, 2002.

[6] F. G. Heineken, H. M. Tsuchiya, R. Aris, On the accuracy of determining rate constants in enzymatic reactions, Mathematical Biosciences 1 (1967) 115–141.

[7] L. A. Segel, On the validity of the steady state assumption of enzyme kinetics, Bulletin of Mathematical Biology 50 (1988) 579–593.

[8] D. R. Hall, N. N. Gorgani, J. G. Altin, D. J. Winzor, Theoretical and experimental considerations of the pseudo-first-order approximation in conventional kinetic analysis of IAsys biosensor data, Analytical Biochemistry 253 (1997) 145–155.

[9] S. Schnell, P. K. Maini, Enzyme kinetics at high enzyme concentration, Bulletin of Mathematical Biology 62 (2000) 483–499.

[10] S. Schnell, C. Mendoza, The condition for pseudo-first-order kinetics in enzymatic reactions is independent of the initial enzyme concentration, Biophysical Chemistry 107 (2004) 165–174.

[11] S. Schnell, Validity of the Michaelis-Menten equation - steady-state or reactant stationary assumption: that is the question, FEBS Journal 281 (2014) 464–472.

[12] S. Schnell, S. M. Hanson, A test for measuring the effects of enzyme inactivation, Biophys. Chem. 125 (2007) 269–274.

[13] S. M. Hanson, S. Schnell, Reactant Stationary Approximation in Enzyme Kinetics, The Journal of Physical Chemistry A 112 (2008) 8654–8658.

[14] M. Whidden, A. Ho, I. M. I., S. Schnell, Competitive reaction mechanisms for the two-step model of protein aggregation, Biophys. Chem. 193–194 (2014) 9–19.

[15] D. A. Lauffenburger, J. Linderman, Receptors: models for binding, trafficking, and signaling, Oxford University Press, 1995.

[16] G. A. Weiland, P. B. Molinoff, Quantitative analysis of drug-receptor interactions: I. Determination of kinetic and equilibrium properties, Life Sciences 29 (1981) 313–330.

[17] M. D. Hollenberg, H. J. Goren, Ligand-Receptor Interactions at the Cell Surface, in: G. Poste, S. T. Crooke (Eds.), Mechanisms of Receptor Regulation, New Horizons in Therapeutics, Springer US, 1985, pp. 323–373.

[18] G. Pettersson, The transient-state kinetics of two-substrate enzyme systems operating by an ordered ternary-complex mechanism, European journal of biochemistry / FEBS 69 (1976) 273–278.

[19] G. Pettersson, A generalized theoretical treatment of the transient-state kinetics of enzymic reaction systems far from equilibrium, Acta Chemica Scandinavica. Series B: Organic Chemistry and Biochemistry 32 (1978) 437–446.

[20] F. Sicilio, M. D. Peterson, Ratio errors in pseudo first order reactions, Journal of Chemical Education 38 (1961) 576.

[21] J. W. Moore, R. G. Pearson, Kinetics and Mechanism, John Wiley & Sons, 1961.

[22] J. F. Corbett, Pseudo first-order kinetics, Journal of Chemical Education 49 (1972) 663.

[23] T. M. Lowry, W. T. John, CCLXIX. Studies of dynamic isomerism. Part XII. The equations for two consecutive unimolecular changes, Journal of the Chemical Society, Transactions 97 (1910) 2634–2645.

[24] C. Castillo-Chavez, B. Song, Dynamical models of tuberculosis and their applications, Mathematical biosciences and engineering: MBE 1 (2004) 361–404.

[25] E. C. Hulme (Ed.), Receptor-ligand interactions: a practical approach, Practical approach series, IRL Press at Oxford University Press, Oxford [England] ; New York, 1992.

[26] C. R. Kahn, Membrane receptors for hormones and neurotransmitters, The Journal of Cell Biology 70 (1976) 261–286.

[27] T. D. Pollard, E. M. De La Cruz, Take advantage of time in your experiments: a guide to simple, informative kinetics assays, Molecular Biology of the Cell 24 (2013) 1103–1110.

[28] L. A. Segel, M. Slemrod, The Quasi-steady State Assumption: A Case Study in Perturbation, SIAM Rev. 31 (1989) 446–477.

[29] T. D. Pollard, A Guide to Simple and Informative Binding Assays, Molecular Biology of the Cell 21 (2010) 4061–4067.

[30] J. Jacquez, The inverse problem for compartmental systems, Mathematics and Computers in Simulation 24 (1982) 452–459.

[31] M. C. Kennedy, A. O'Hagan, Bayesian calibration of computer models, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63 (2001) 425–464.

[32] H. W. Engl, C. Flamm, P. Kgler, J. Lu, S. Mller, P. Schuster, Inverse problems in systems biology, Inverse Problems 25 (2009) 123014.